Update on the CMIP5 Observation Data Base

Jerry Potter NASA GSFC

Information furnished by:

LLNL

Dean Williams
Charles Doutriaux
Renata McCoy
Bob Drach
Mike Ganzberger

JPL

Robert Ferraro Duane Waliser Joao Teixeira Luca Cinquini

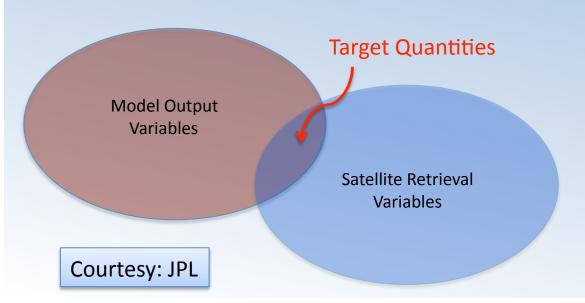
Outline

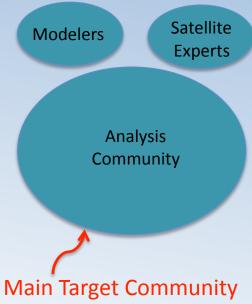
- Background
- What has been agreed upon
- What remains
- Publishing data
- Browsing data
- Deadlines

Satellite Observations for CMIP5 Simulations Some Basic Tenets of this Activity

- •To provide the community of researchers that will access and evaluate the CMIP5 model results access to *analogous sets* (in terms periods, variables, temporal/spatial frequency, dissemination) of satellite data.
- To be carried out in close coordination with the corresponding CMIP5 modeling entities and activities — Working Group on Coupled Modeling (WGCM).

 To directly engage the observational (e.g. mission and instrument) science teams to facilitate production of the corresponding data sets and documentation.





Satellite Observations for CMIP5 Simulations Main Tasks for CMIP5

- 1. Use the CMIP5 simulation protocol as guideline for deciding which observations to stage in parallel to model simulations.

 Target is monthly averaged (CMON, AMON) products on 1 x 1 degree grid
- 2. Work with satellite community to identify data sets [e.g., AIRS, MLS, TES, QuikSCAT, CloudSat, Topex/Jason, CERES, TRMM, AMSR-E, TRMM]
- 3. Work with observational teams to produce 4-5 page Technical Note describing strengths/weaknesses, uncertainties, dos/don'ts regarding interpretations comparisons with models. (at graduate student level)
- Transform satellite observations into CMIP5 compliant format (to the greatest extent possible should look like model output)
 Not new products, but reformatted existing products, with perhaps some interpolation
- 5. Provide a strategy for accessing them that has close parallels to the model data archive (using ESG).
- 6. Advertise availability of observations for use in CMIP5 analysis.

Courtesy: JPL

NASA Recommended Datasets for CMIP5

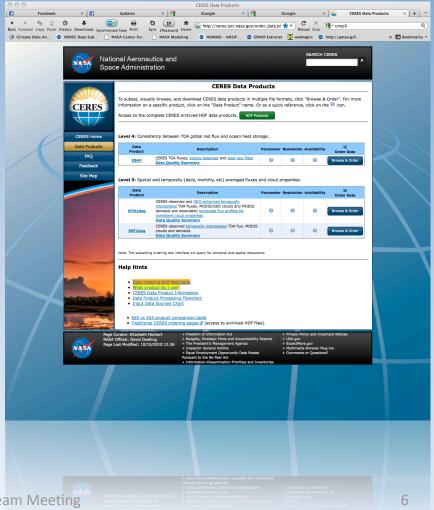
Match up of available NASA datasets to WGCM priority list

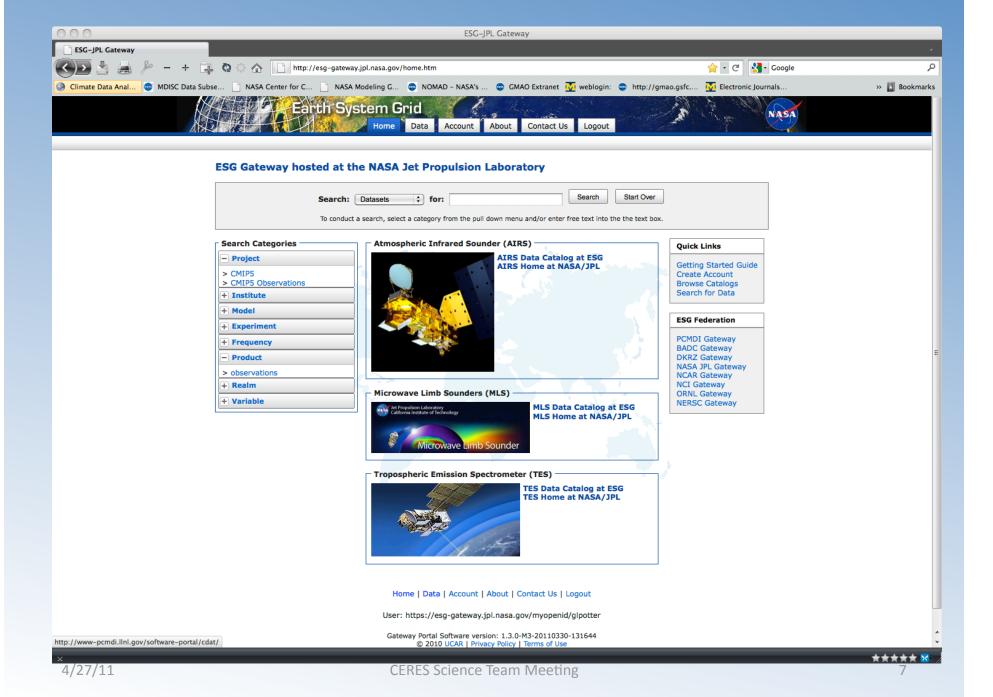
Model	Dataset	Time Period	Comments
Atm Temperature (200,850hPa)	AIRS (≥ 300 hPa) MLS (< 300 hPa)	9/02 – 8/04 -	AIRS +MLS needed to cover all pressure levels
Zonal and meridional wind (200,850 hPa)	No obvious match		Reanalysis is the best product
Specific humidity (200, 850 hPa)	AIRS (≥ 300 hPa) MLS (< 300 hPa)	9/02 – 8/04 -	AIRS +MLS needed to cover all pressure levels
Sea level pressure	No obvious match		Reanalysis is probably the best product match
Surface (10m) zonal and meridional wind	QuikSCAT CCMP	1999 – 2009 7/87 – 12/09	Oceans only. No land products. CCMP is a multi-sensor variational analysis product
Ocean surface zonal and meridional wind stress	QuikSCAT CCMP	1999 – 2009 7/87 – 12/09	Oceans only. No land products. CCMP is a multi-sensor variational analysis product
Sea surface temperature	AMSR-E	6/02 -	SST science team recommends multiple products
TOA reflected shortwave radiation and OLR	CERES	3/00 -	
TOA longwave and shortwave TOA clear-sky fluxes	CERES	3/00 -	
Total precipitation	TRMM GPCP	1997 - 2/79 - 4/08	GPCP is an analysis product
Cloud cover	MODIS	2/00 -	
Precipitable water	SSM/I	7/87 -	
Sea surface height	TOPEX/JASON series	10/92 -	Project scientist recommends converting the AVISO product
PL ice	NSIDC CERES Science Tea	m Meeting	microwave product would be best. More investigation is needed.

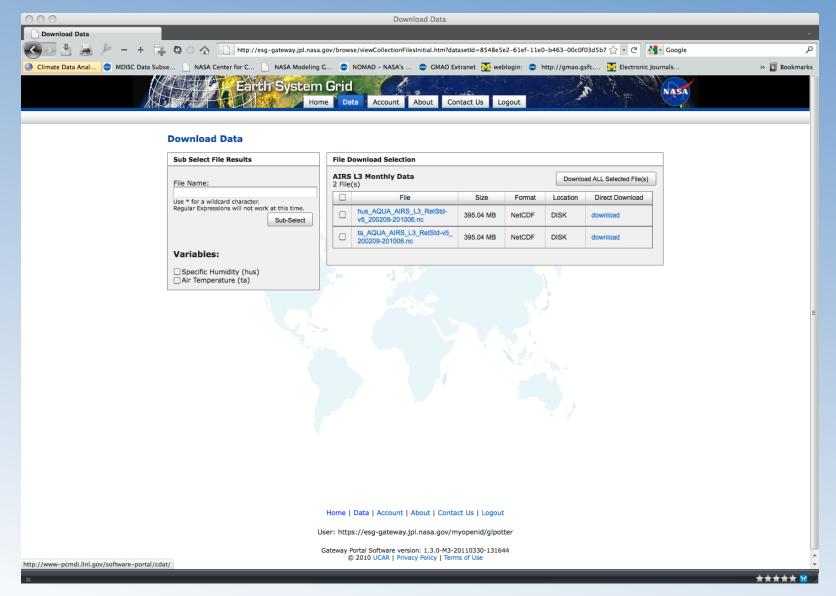
Courtesy: JPL

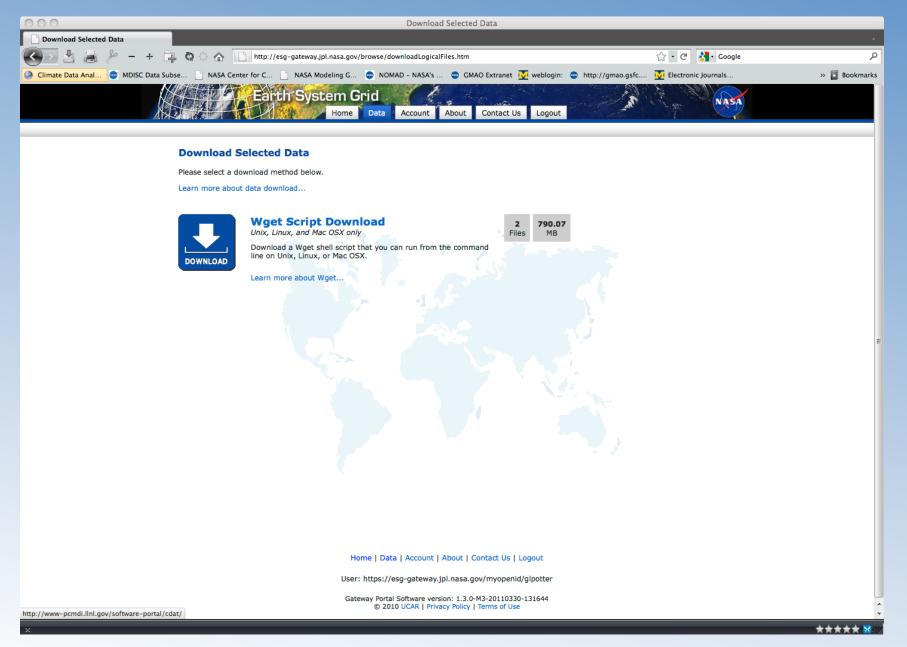
Why use the Earth System Grid (ESG) to distribute observations for CMIP5?

- The CERES data Product site
 - Works great
 - Easy access
 - But, doesn't appear to users in the same way as the CMIP5 data
- CERES, along with other satellite data products will reside along with the CMIP5 model output products









Satellite Observations for CMIP5 Simulations Example Technical Note – AIRS

Atmospheric Infrared Sounder/Advance Microwave Sounding Unit (AIRS/AMSU)

Air Temperature Description

1 Intent of This Document

1a) This document is intended for users who wish to compare satellite derived observations with climate model output in the context of the CMIP5/IPCC historical experiments. Users are not expected to be experts in satellite derived Earth system observational data. This document summarizes essential information needed for comparing this dataset to climate model output. References are provided at the end of this document to additional information for the expert user.

This NASA dataset is provided as part of an experimental activity to increase the usability of NASA satellite observational data for the model and model analysis communities. This is not a standard NASA satellite instrument product. It may have been reprocessed, reformatted, or created solely for comparisons with the CMIP5 model. Community feedback to improve and validate the dataset for modeling usage is appreciated.

Dataset File Name (as it appears on the ESG): ta_Amon_AIRS_amip_rlilpl_200209-201006.nc

1b) Technical point of contact for this dataset: Baijun Tian, Baijun Tian@jpl.nasa.gov

2. Data Field

This data product is a regularly gridded, monthly averaged air temperature measured by AIRS during 2002-2010.

CF variable name, units: ta, K.

Vertical resolution: CMIP5 mandatory levels

Horizontal resolution: 1 degree.

Temporal resolution and extent: monthly averaged between September 2002 and June 2010.

Coverage: Global

The vertical pressure levels (plev) include all the CMIP5 mandatory levels from 1000 hPa to 10 hPa. However, we only provide the valid data up to 300 hPa and assign a missing value (1e-20) for levels above 300 hPa because AIRS measurements are not as reliable for levels above 300 hPa as other instruments such as Microwave Limb Sounder (MLS), which is specially designed for the accurate measurements of the atmospheric profiles in the upper troposphere and lower stratosphere. The MLS measurements for levels above 300 hPa are provided as a separate dataset. It is therefore highly recommended for a user to combine the AIRS and MLS datasets to create a complete temperature profile from 1000 hPa to 10 hPa.

3. Data Origin

The data used to make this product was obtained from the Goddard Earth Science (GES) DISC data access [1].

The AIRS/AMSU instrument suite is carried on the NASA Aqua spacecraft, in a sunsynchronous orbit at 1:30 local time. The southward/northward moving observations are
obtained during daytime/nighttime. (See Section 6 below for an Overview of the AIRS/AMSU
instrument suite.) The AIRS/AMSU air temperature is derived from infrared and microwave
radiances measured from space, so is not an in situ measurement. The infrared emission
radiations emitted by different Earth scenes are remotely sensed by a spectrometer, and the
microwave observations are obtained by a radiometer [2]. Among the 2378 infrared spectral
channels, 147 channels are used in the first temperature profile retrieval; an additional seven
channels are used in the second temperature profile retrieval; 12 or 15 microwave channels are
utilized for temperature sounding [3]. First, measurements are transformed into calibrated
radiances for all footprints and all channels [4]. Then, physical quantities such as the air
temperature are derived ("retrieved") from these geolocated radiance products [5]. The retrieved
physical quantities are then averaged over a month [6]. The data we obtained from the GES
DISC [1] was at this last processing level. We then applied an additional processing step to adapt
the data according to the CMIP5 model output format.

The values described here are means of the daytime and nighttime values, provided there are enough observations in each category to make the values statistically significant. The minimum is 20 observations each, except for latitudes beyond +/- 80 degrees, where we relax the limits to compensate for a much lower number of observations. Since clouds have a significant effect on observed infrared radiances (see section 4-1 below), the retrieval process includes steps to retrieve the temperature from radiance in the presence of clouds [3]. The horizontal resolution of each AIRS/AMSU scene is 45 km, and the instrument sample in a swath 30 scenes wide (see Figure 3 below), yielding 324,000 scenes per day. However, the atmospheric temperature can be inferred in about 70% of these scenes, with the remainder affected by thick clouds or precipitation.

This data product is the monthly average of the AIRS/AMSU retrieved temperature profiles in the regularly gridded 1 degree by 1 degree latitude and longitude boxes. The AIRS/AMSU temperature profile is originally in 28 pressure levels that are the superset of the CMIP5 mandatory levels. The temperature values at only the CMIP5 levels are included in this data product.

2

Satellite Observations for CMIP5 Simulations Example Technical Note – AIRS

4 Validation

AIRS retrievals have been validated against a variety of in situ data (radiosondes, ship-based measurements), other remote measurements, from other satellites, and model-generated data. The table below summarizes these findings and can be found in reference [7].

Geophysical Conditions Studied	Uncertainty Estimate
Ocean, surface to lower stratosphere	1K
Non-polar land, 1-2 km to lower stratosphere	1K
Non-polar land, surface to 1-2 km	1-2 K
Polar land.	1-2 K

Table 1: uncertainty estimate for different conditions.

The uncertainty estimates are calculated based on the difference between AIRS retrievals and radiosonde observations and are valid in the troposphere, between the boundary layer and 300 mbar.

5. Consideration for Model-Observation Comparisons

Because this data product is observational data, there are several aspects that distinguish this product from model outputs. The user of this data product should be aware of them in order to make judicious model-observation comparisons.

5.1 Clouds influence

AIRS/AMSU coverage is limited by the presence of optically thick clouds because AIRS is an infrared instrument. The combination of infrared and microwave radiances allows retrieval of high-resolution temperature profiles for infrared cloud fraction (the product of emissivity and coverage) up to about 70% [3]. While AMSU retrievals based purely on microwave observations are currently available, they are not included with the AIRS/AMSU retrievals because of significantly lower vertical resolution. This limitation of the infrared measurement makes the AIRS/AMSU observation scene dependent and in turn, causes a spatially inhomogeneous sampling as illustrated on Figure 1. The AIRS sampling is low (~40) in cloudy regions, such as the Intertropical Convergence Zone (ITCZ) (e.g., the equatorial western Pacific warm pool) and the midlatitude storm tracks (e.g., north Pacific, north Atlantic and 60S latitude belt). The AIRS sampling is high (~150) in clear regions, such as subtropics and midlititude land regions. See reference [8] for more on the implication of cloud-induced sampling in AIRS/AMSU observations.

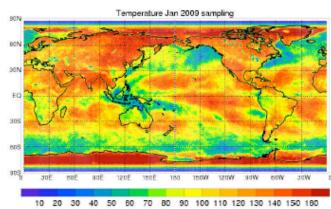


Figure 1: Temperature sampling at 500 hPa for the month of January 2009.

5.2 Temperature bias trend

There is a spurious trend in AIRS temperature retrievals, giving an apparent cooling of the lower troposphere of ~100 mK/year. This is suspected to come from incorrect handling of rising atmospheric CO2 levels [9], and should be reduced in the next release of AIRS data.

5.3 Asynoptic Time Sampling

Because Aqua satellite with a sun-synchronous polar orbit, it samples at two fixed local solar times (e.g. 1:30 AM and 1:30 PM at the equator) so cannot resolve the diurnal cycle. AIRS observes at a given latitude at the same local time (to within several minutes). In contrast, typical model monthly averaged outputs contain the averaged values over a time series of data within a fixed time interval (e.g. every 6 hours). For temperature in the upper atmosphere, this difference is not likely a problem although for regions influenced by deep convection and its modulation of the diurnal cycle (e.g. tropical land masses), this time sampling bias should be considered.

5.4 Inhomogeneous Sampling

Because the monthly averaged value in this AIRS data product is an average over observational data available in a given grid cell (see Figure 1), the number of samples used for averaging varies with the geo-location of the cell. Because of the convergence of longitude lines near the poles, the time range of data collection broadens as one moves from the equator toward either pole, with the ranges in the polar regions including all times of day and night [10]. So, there are more observations in the regions near the poles (~70° to ~85°) than the rest of the area.

6. Instrument Overview

Satellite Observations for CMIP5 Simulations Example Technical Note – AIRS

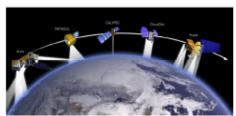


Figure 2: NASA's A-train group of Earth observing satellites.

Launched into Earth-orbit on May 4, 2002, the Atmospheric Infrared Sounder, AIRS, is one of six instruments on board the Aqua satellite, part of the NASA Earth Observing System AIRS along with its partner microwave instrument, Advanced Microwave Sounding Unit (AMSU-A), observe the global water and energy cycles, climate variation and

trends, and the response of the climate system to increased greenhouse gases. The term "sounder" in the instrument's name refers to the fact that temperature and water vapor are measured as functions of height.

AIRS and AMSU-A share the Aqua satellite with the Moderate Resolution Imaging Spectroradiometer (MODIS), Clouds and the Earth's Radiant Energy System (CERES), and the Advanced Microwave Scanning Radiometer-EOS (AMSR-E). Aqua is part of NASA's "A-train" satellite constellation (see Figure 2), a series of high-inclination, sun-synchronous satellites in low Earth orbit designed to make long-term global observations of the land surface, biosphere, solid Earth, atmosphere, and oceans.

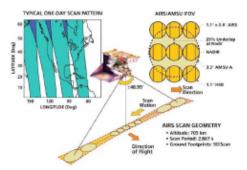


Figure 3: AIRS scanning and coverage geometry.

AIRS coverage is pole-to-pole and covers the globe two times a day. Because the swaths (scanning sweeps) do not overlap at low latitudes, some points near the equator are missed. However, these points are eventually scanned within 2-3 days. As depicted on Figure 3, AIRS scans laterally with respect to its direction of flight. With the scanning angle being 49.5 degree about nadir, the swath width is 1650 km. One orbit period is 98.8 minutes [11].

7 References

- [1] http://disc.sci.gsfc.nasa.gov/AIRS/data-holdings
- [2] Hartmut H. Aumann et al., "AIRS/AMSU/HSB on the Aqua Mission: Design, Science Objectives, Data Products, and Processing Systems", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 41, NO. 2, FEBRUARY 2003.
- [3] Joel Susskind et al., "Retrieval of Atmospheric and Surface Parameters From AIRS/AMSU/HSB Data in the Presence of Clouds", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 41, NO. 2, FEBRUARY 2003, page 390.
- [4] Level-1B AIRS IR
- [5] Level-2 Standard Products Ouick Start
- [6] Level-3 Standard 1x1° Gridded Products Quick Start
- [7] V5 CalVal Status Summary.pdf, p8.
- [8] Fetzer, E. J., et al., (2006), Biases in total precipitable water vapor climatologies from Atmospheric Infrared Sounder and Advanced Microwave Scanning Radiometer, J. Geophys. Res., 111, D09S16, doi:10.1029/2005/D006598.
- [9] Murty G. Divakarla et al., "Validation of Atmospheric Infrared Sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts", JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 111, D09S15, doi:10.1029/2005/D006116, 2006, page 18.
- [10] Claire L. Parkinson, "Aqua: An Earth-Observing Satellite Mission to Examine Water and Other Climate Variables", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 41, NO. 2, FEBRUARY 2003.
- [11] http://airs.jpl.nasa.gov/instrument/coverage/

8. Revision History

Rev 0 - 03/07/2011

Satellite Observations for Climate Modeling SUMMARY

 Pilot Project to establish a NASA-wide capability for the climate modeling community to support model-to-data intercomparison. This involves IT, satellite retrieval, data set, modeling and science expertise.

Longer Term

 Provide a science "bridge" between models and satellite observations to facilitate model improvement and reduce projection uncertainty. This is also a focus of the new JPL Center for Climate Science.

Ultimately

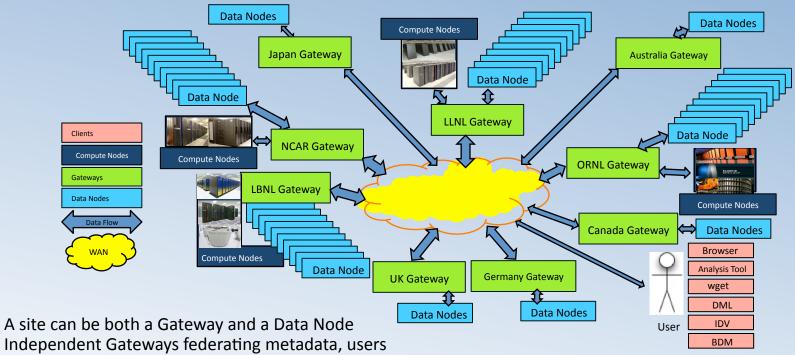
 Utilize feedback/community collaboration to develop future climate-critical satellite missions. The modeling community has yet to be galvanized to provide feedback to the satellite-development community.

This project is on course to deliver NASA satellite data sets by the end of May for the evaluation of CMIP5 climate model archive and impact the IPCC AR5.

What is the Earth System Grid?

- Gateway Nodes: where data is discovered and requested
- Gateway Node functionality includes
 - portals, registration and user management
 - search capability, distributed metadata
 - may be customized to an institution's requirements, topical focus

- Data Nodes: where data is actually stored or archived
- Data Node functionality includes
 - data publication (making it visible to an ESG Gateway)
 - data reduction/analysis support
 - possible minimalist deployment without services
 - delivery services to ESG end users



- Any user can discover any data from any Gateway
- Each data node publishes to one or more Gateways
- Specific data collections are managed through specific Gateways no

Courtesy: Dean Williams, LLNL

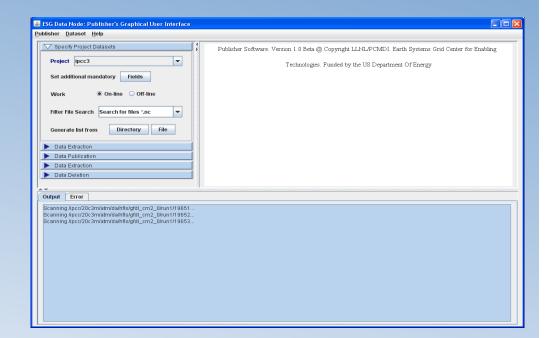
Critical Functions

- **Data Versioning** A user must be able to uniquely identify, access, and cite any published object (e.g., files, datasets, collections, and aggregations) in the database as it existed at any point in the lifetime of the database.
- **Replication** A node can choose to replicate a collection of the datasets published by a different node. This includes replicas of aggregated datasets.
- Metrics Quantitative measurement and reporting of the usage and performance of the ESG Enterprise system.
- **Bulk Data Movement (BDM)** The transfer of large collections between sites reliably and with good performance and with a high-level of easy-of-use.
- **Provenance** important for science because it helps to interpret and reproduce the results of an experiment; to understand the chain of reasoning used in the production of a result; to verify that the experiment was performed according to acceptable procedures, to track who performed the experiment and who is responsible for its results.

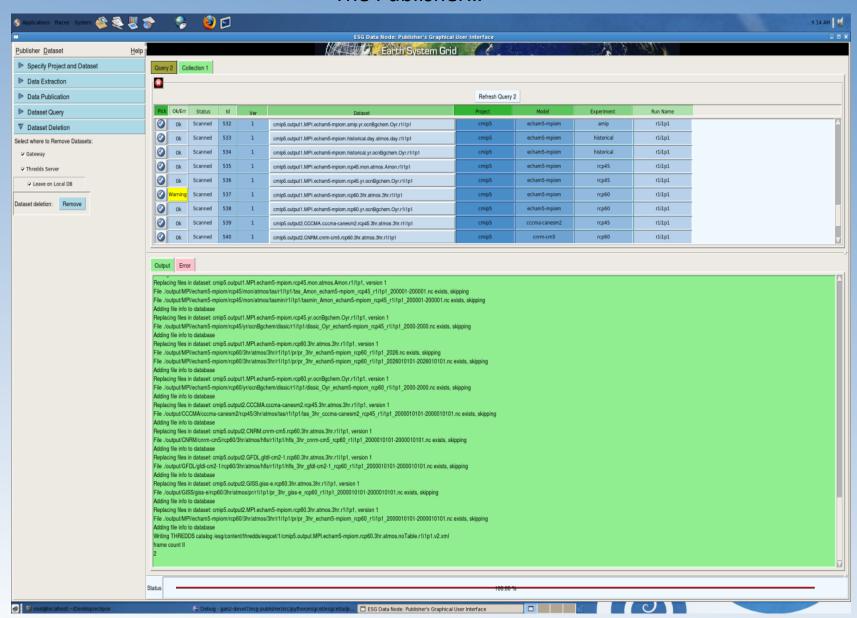
Once the data is rewritten to have the same "look" as model data the meta data is published at the gateway

Tools for publishing (sending the meta-data to the gateway)

- Versioning is maintained
- Publishing, unpublishing, help via web browser and comments collected and displayed.
- The Delete options allow a user to remove datasets from the Thredds Server, Gateway with the option to keep it on the local DB.



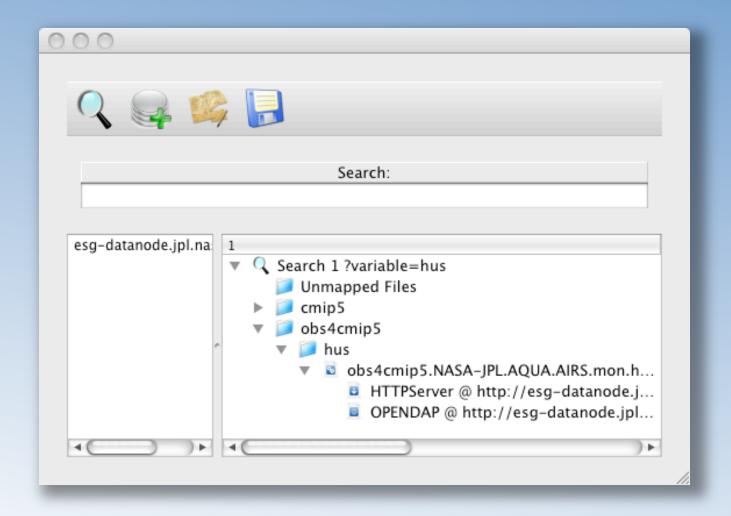
The Publisher...



ESG Search Capabilities

- User Registration.
- User and Group Management
- Login
- Data Browsing
 - Support for file system-like hierarchies and high-level associative arrangements
- Data Search
 - The search component provides simple and familiar text-based search as well as advanced capabilities using a dynamic faceted query interface for detailed metadata inquiry
- Data Download
- Data Transfer
- Data Visualization and Sub-Setting
- Model Metadata Trackback
- User Workspace
- Metrics Reporting

ESG Search capabilities



Deadlines

- CMIP5
- Observations for CMIP5
 - Plan for end of May, 2011
 - May slip

Object Based Evaluation of GCM-Simulated Clouds and Radiation During the 1998 El Nino- La Nina Transition

Derek J. Posselt¹, Chuan-Yuan Hsu¹, Jerry Potter¹, and Andrew Jongeward²

¹University of Michigan ¹University of Maryland





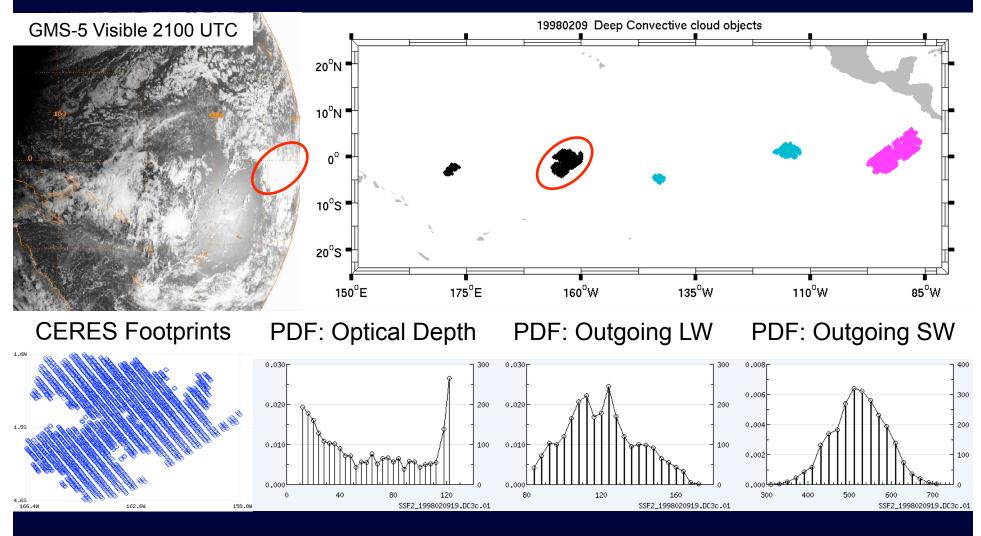
Introduction and Motivation

- Importance of realistically simulating convective cloud systems in climate models
 - Radiation budget
 - Precipitation
 - Large-scale circulation
- There is a strong coupling between convection, radiation, and large scale circulation in the real Earth system
- Challenges:
 - Improvements in models increasingly need to be process-based
 - Desire to exploit the tremendous amount of detailed information contained in current and next generation satellite systems
 - How to evaluate simulated cloud systems in GCMs, which are not expected to reproduce features at the observed time and location

Objectives

- 1. Explore new ways of comparing/merging models and observations
- 2. Evaluate simulated deep convective cloud systems in the tropics using multisensor observations
- 3. Examine sources of error inherent in the model evaluation process
- Use 1997-1998 El Nino as a test case
 - Force model in AMIP mode, examine response of cloud systems
 - Is convection properly coupled to the large scale?

Cloud Objects 9 February 1998



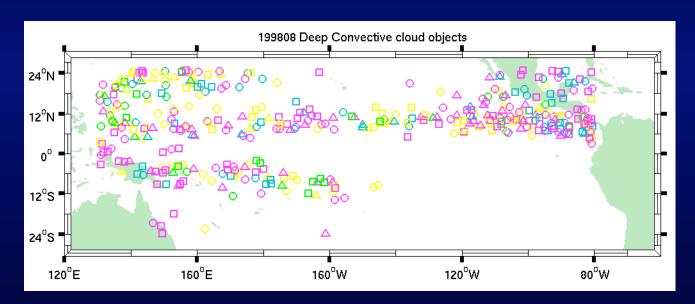
Deep Convective Cloud Observations During the 1997/1998 El Nino

- Deep convective cloud objects identified from CERES instrument on TRMM (Xu et al. 2005, J. Climate)
- Defined as a contiguous region of cloud between [-25,25] deg. latitude and [130,-80] deg. Longitude with:
 - Optical depth > 10
 - Cloud top height > 10 km
- Obtained for the time period of CERES on TRMM (January – August 1998, March 2000)
- Examples follow

All Objects Grouped by Month

- Square: < 150 km</p>
- Circle: 150-300
- Triangle: 300+ km

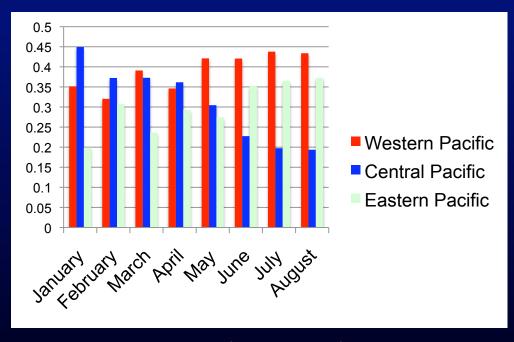
- Green: 0600 0900 LST
- Yellow: 0900 1200 LST
- Magenta: 1200 1500 LST
- Blue/Green: 1500 1800 LST



Signal of the 1997/1998 El Nino in Deep Convective Cloud Objects

- Examine the fraction of objects occurring in the
 - West: [130,180] deg. lon
 - Central: [-180,-130] deg. lon
 - East: [-130,-80] deg. lon

 Systematic decrease in fraction of objects occurring in Central Pacific



Numerical Model

- NASA GEOS-5 model run in two configurations
- Reanalysis (MERRA) 0.5 x 0.625 deg. lat/lon grid
 - Flow, thermal and moisture fields, and precipitation are constrained by observations.
 - Cloud properties and convection generated by model physics, but cloud systems should occur at ~ the correct place and time.
- Prescribed SST (AMIP) 2.0 x 2.5 deg. lat/lon grid
 - Prescribed ocean state, but atmosphere is free to evolve.
 - No constraint to place clouds in the correct place and time
- Resolution far too coarse to resolve mesoscale circulations—convection is entirely parameterized
- Desire realistic interaction between convection, radiation, and large scale circulation

Comparison Methodologies

- 1. Compare objects with MERRA grid boxes collocated in space and time
- 2. Evaluate free-running GEOS-5 by
 - Identifying regions where we would expect persistent deep convection (consistent with the large scale circulation)
 - Performing the comparison only in these regions.

Cloud Objects

- Database of cloud properties subset into different cloud types
- Contain information on cloud physical and radiative properties
- Can be treated as a sample of the variety of realworld deep convective systems
- PDFs contain information on the relative frequency of thin vs thick cloud and on the distribution of values of TOA radiative flux
- Challenge: extracting similar PDFs from a model

Mapping Model to Observation Space

- Observations consist of a sample of CERES measured and retrieved quantities on CERES 10x10 km footprints
- Need to disaggregate model grid boxes into subgrid scales
- Methodology
 - Run ISCCP subcolumn generator (max-random overlap)
 - Divide cloud mass among subcolumns and run Fu-Liou radiative transfer model
 - Compare PDFs of cloud optical depth, total water path, and Outgoing LW and SW radiative flux.